

Bounds to Truncation Errors in Biorthogonal Polynomial Approximations, With Illustrative Applications to Gamma Ray Transport Distributions*

L. V. Spencer

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(January 29, 1971)

Problems of gamma ray and neutron penetration in infinite, homogeneous media can be solved in terms of infinite sums of a certain type of biorthogonal polynomial system. Many calculations have used a few terms, perhaps 4 to 8, successfully due to good apparent convergence. This paper develops expressions for bounds to the truncation error of these finite sums. Examples and applications are given of the functions used to obtain the error bounds.

Key words: Biorthogonal polynomials; gamma ray transport; moment method; neutron transport; polynomial truncation errors; Sievert function.

1. Introduction

There is an important class of radiation transport problems for which it is feasible to calculate spatial moments

$$M_n = \frac{1}{n!} \int_0^\infty dz z^n F(z), \quad (1)$$

or

$$M_n = \frac{1}{n!} \int_{-\infty}^\infty dz z^n F(z), \quad (1')$$

by numerical integration of a recursive system of Volterra integral equations of the second kind [1, 2].¹ A detailed description of the procedure for obtaining and evaluating these equations is not pertinent to the discussion here; but it is important that accurate calculation is feasible only for a limited number of the moments, say 6 to 10.

In some of these problems every other moment is unobtainable because the recursive system does not reach it; in other problems every other moment vanishes; and in still other problems it turns out to be advantageous to distinguish between even and odd component distributions, both of which utilize alternate moments only. Thus one continually runs into the problem of constructing distributions which depend on a space variable z , while utilizing powers of the variable z^2 through the use of alternate moments.

There is by now a sizeable body of literature on this subject which references [1-13] do not exhaust. Two general approaches to the problem of constructing the distributions have been used,

AMS Subject Classification: Primary 4130, Secondary 6580.

*Work supported by funds from the Office of Naval Research and the Office of Civil Defense.

¹ Figures in brackets indicate the literature references at the end of this paper.

namely with polynomials and with sums of functions such as exponentials [2]. The latter method has been referred to as "function-fitting," and a variety of algorithms for determining coefficients in such representations have been used.

The problem which has concerned us is the following: Data from "moment" calculations exists, has proven very useful [5-7], and has even been used as a kind of standard against which data from other procedures with known and possibly severe built-in limitations has been compared [8-10]. Yet there has been no way to assess the errors in the data constructed from the moments. Hence the accuracy of these data, although apparently high and generally assumed so, is not firmly established.

This paper is the second in a series which attempts to put such data on a firmer foundation. There appear to be at least four questions which call for study: Let N be the number of moments which an approximation utilizes. Can one be certain of convergence in the limit $N \rightarrow \infty$, (1) for polynomial approximations or (2) for specific function fitting algorithms? Likewise, what is the "truncation error" due to use of only a finite set of moments, (3) in polynomial approximations, or (4) in specific function-fitting approximations?

The first paper in this series, reference [4], was a study of the polynomial system, which essentially answered question (1). A few comments about this will be made in the next section. In this second paper we discuss procedures for estimating the truncation error for the polynomial approximations, i.e., question (3). Two approaches are given, of which one is more suited to the radiation transport problems which are our main concern.

We also have ways of getting at answers to both of the questions (2) and (4), and we hope to make these the subject of one or more additional papers. In fact, function-fitting procedures can be more flexible and accurate than polynomial methods. But problems of convergence and estimation of truncation errors for function-fitting approximations are intimately bound up with the properties of related orthogonal functions, and that is the reason for delaying study of function-fitting approximations until the polynomial problems have been considered.

In the next section we identify the pertinent polynomials by giving some of their main properties. Reference should be made to references [1-4], particularly reference [4], for a more complete discussion of the properties reviewed here.

2. U_n^k Polynomial Systems and Approximations

The polynomial representations to which we have referred are usually written

$$F(z) = \sum_{n=0}^{\infty} F_n U_n^k(z), \quad (2)$$

and our interest also lies in truncated approximations to the infinite series. The independent variable in (2) is written z , with $0 \leq z < \infty$; but it may also be $|z|$, with $-\infty < z < \infty$. The U_n^k are biorthogonal to an adjoint set of polynomials which we designate \hat{U}_n^k ; and the weight function for the system is $z^k e^{-z}$:

$$\int_0^{\infty} dz z^k e^{-z} \hat{U}_n^k(z) U_m^k(z) = \delta_{nm}, \quad (3)$$

where δ_{nm} is Kronecker's delta. The adjoint polynomials, \hat{U}_n^k , are functions of z^2 rather than z or $|z|$. The two sets, U_n^k and \hat{U}_n^k , satisfy adjoint *third* order differential equations, and have been shown to be the only biorthogonal polynomials obeying differential equations of this order [3]. We also note that generating functions for the U_n^k are elementary [4],

$$(\sqrt{1-q})^{-k-1} \exp(z - z/\sqrt{1-q}) = \sum_{n=0}^{\infty} q^n U_n^k(z); \quad (4)$$

and that while generating functions for the \hat{U}_n^k can likewise be written in closed form, only the cases $k=0, 1$ are as simple as (4) [4]:

$$(1-q)^{-1} \cos [z\sqrt{q/(1-q)}] = \sum_{n=0}^{\infty} q^n \hat{U}_n^0(z),$$

$$[q(1-q)]^{-1/2} \sin [z\sqrt{q/(1-q)}] = \sum_{n=0}^{\infty} q^n \hat{U}_n^1(z). \quad (4')$$

Generally speaking, convergence of the series in eq (2) appears to be rapid enough for many important practical problems. Typically the first four or five terms seem to give impressive accuracy over a wide range of z values, say $0 \leq z \leq 15$.

The Fourier transforms of the U_n^k and the Laplace transforms of the \hat{U}_n^k are simple power functions in the complex plane and have biorthogonality properties on an appropriate contour. From these functions one can show that convergence of the series (1) depends on the location of the singularities of the Fourier transform of $F(|z|)$ [4]. In the gamma ray and neutron problems of interest, the form of the transport equation leads one to expect the singularities to lie on the real axis to the right of a known point. Convergence is then expected if very simple rules are obeyed in choosing a scale parameter (for z) so that all singularities fall in a hyperbolic-shaped region of the complex plane [4].

This leaves unanswered the question of the size of the truncation error which results from use of a finite rather than infinite sum. Our approach here is to obtain "error limit functions" which enable us to bound the error for all values of z . Our method for developing these functions is simple and not particularly new. Nevertheless, because neither the study of truncation errors nor the use of biorthogonal systems is familiar to many physicists and engineers, we put the problem in a fairly general context and in the process include older material to make the discussion more complete.

3. The Adjoint Function and the Norm

One can often, though not always, construct an adjoint function to $F(z)$ which we designate $\hat{F}(z)$,

$$\hat{F}(z) = \sum_{n=0}^{\infty} F_n \hat{U}_n^k(z). \quad (5)$$

In addition, the number $\|F\|^2$,

$$\|F\|^2 = \int_0^{\infty} dz z^k e^{-z} \hat{F}^*(z) F(z) = \sum_{n=0}^{\infty} |F_n|^2, \quad (6)$$

has the properties of a norm. Since these series may or may not converge depending on the sequence of F_n values, these concepts are useful mainly when (6) converges so that the function has a *finite* norm of this type.

In this connection we can bring out some important ideas by rewriting (2), (5), and (6) with an arbitrary renormalization of the polynomials:

$$F(z) = \sum_{n=0}^{\infty} (c_n F_n) \{ c_n^{-1} U_n^k(z) \}, \quad (2')$$

$$\hat{F}(z) = \sum_{n=0}^{\infty} (c_n F_n) \{ c_n \hat{U}_n^k(z) \}, \quad (5')$$

² The asterisk in (6) indicates complex conjugate.

where the subscript indicates a renormalization, and

$$\|F\|_r^2 = \sum_{n=0}^{\infty} |c_n F_n|^2. \quad (6')$$

A comparison with eq (2) makes it clear that the biorthogonality relation holds for the renormalized polynomials (U_n^k/c_n) because the c_n, c_m in (2) can be canceled for all n and m . Further, the distribution $F(z)$ is not changed by this renormalization. But the adjoint function and the norm *are* changed, hence the use of the subscript r in (5') and (6'). If the constants c_n decrease with n , it is possible for $\|F\|_r$ to be finite when $\|F\|$ is not. Correspondingly, the series for $\hat{F}_r(z)$ may converge where $\hat{F}(z)$ does not.

Before proceeding further, we should note that this discussion, and the discussion to follow, applies to orthogonal as well as biorthogonal systems. But the indicated renormalization generates a difference between two sets of functions which would otherwise be identical for the orthogonal case. Hence application of the argument to orthogonal systems involves treating them as two distinct biorthogonal sets.

4. Transformation Kernels

We need to refer to one other type of function before turning our attention to the problem of estimating limits for the error of truncated series. We denote this bivariate function by $T^k(y, z)$,

$$T^k(y, z) = \sum_{n=0}^{\infty} \{c_n^{-1} U_n^k(y)\} \{c_n^{-1} U_n^k(z)\}. \quad (7)$$

This series may diverge. But if integrals over $T^k(y, z)$ converge, one can assign to this function the properties

$$\{c_n^{-1} U_n^k(z)\} = \int_0^{\infty} dy y^k e^{-y} T^k(y, z) \{c_n \hat{U}_n^k\} \quad (8)$$

and

$$F(z) = \int_0^{\infty} dy y^k e^{-y} T^k(y, z) \hat{F}_r(y). \quad (8')$$

Hence we refer to T^k as a transformation kernel for the system.

The diagonal case, $T^k(z, z)$, is more useful to us than $T^k(y, z)$, and we give this one-variable function a special notation,

$$[D^k(z)]^2 = T^k(z, z) = \sum_{n=0}^{\infty} \{c_n^{-1} U_n^k(z)\}^2. \quad (9)$$

If this function exists, it is clearly positive; and it is smaller the larger the c_n 's.

5. Some Special Properties and Cases

The normalization for the U_n^k polynomials was chosen fortuitously in reference [1]. One can show that for $c_n = 1$, all the $T^k(y, z)$ can be expressed in terms of modified Bessel functions [4]; while for other choices of c_n such simple expressions probably don't exist. Specifically, for $k=0$,

$$T^0(y, z) = (2/\pi) K_0(\sqrt{y^2 + z^2}) e^{(y+z)}, \quad (10)$$

and

$$[D^0(z)]^2 = e^{2z}(2/\pi)K_0(z\sqrt{2}). \quad (11)$$

Since K_0 diverges logarithmically as its argument tends to zero, the series expression for these functions converges except for $y=z=0$. Divergence at this point also follows from the asymptotic expression for $U_n^k(0)$: for $n \rightarrow \infty$ [4],

$$U_n^k(0) \rightarrow A_k \sqrt{n+1}^{(k-1)}, \quad (12)$$

with $A_0 = 1/\sqrt{\pi}$. Further, one expects that the series (2) will converge for all values of z if

$$\sum_{n=0}^{\infty} \frac{1}{\sqrt{n+1}} |F_n| < \infty. \quad (12')$$

Since this condition is barely more restrictive than the condition that the norm $\|F\|$ be finite, we expect that for $k=0$, the series (2) usually converges for all values of z when the norm defined by (6) is finite.

For $k \geq 1$, the T^k and D^k are more complicated:

$$\begin{aligned} [D^1(z)]^2 &= \frac{1}{\pi} K_2(z) e^{2z}, \\ [D^2(z)]^2 &= \frac{1}{\pi} \left\{ \frac{1}{2} K_4 - \frac{\sqrt{2}}{z} K_3 + \frac{1}{z^2} K_2 \right\} e^{2z}, \\ [D^3(z)]^2 &= \frac{1}{\pi} \left\{ \frac{1}{4} K_6 - \frac{3/\sqrt{2}}{z} K_5 + \frac{9/2}{z^2} K_4 \right\} e^{2z}, \\ [D^4(z)]^2 &= \frac{1}{\pi} \left\{ \frac{1}{8} K_8 - \frac{3/\sqrt{2}}{z} K_7 + \frac{21/2}{z^2} K_6 - \frac{9\sqrt{2}}{z^3} K_5 + \frac{9/2}{z^4} K_4 \right\} e^{2z} \\ T^k &= e^{(y+z)} \frac{\partial^2}{\partial y \partial z} \{ e^{-(y+z)} T^{k-1} \}. \end{aligned} \quad (13)$$

Divergence for $z \rightarrow 0$ occurs for all k , and becomes progressively stronger as k increases. In fact, to guarantee convergence at $z=0$ in (7), it is necessary to choose c_n values which tend to ∞ as

$$c_n = \sqrt{n+1}^{k+\epsilon}, \quad \epsilon > 0. \quad (14)$$

This follows from eq (12). Basically, this factor compensates for weight functions which include the factor z^k , a factor diminishing the amount of information about z near zero which the moments contain. In any case, use of (14) or a different choice which similarly assigns values other than unity to the c_n means that T^k and D^k are not obtainable in such an attractive form as (10) and (13). Such choices also place greater demands on the convergence of the F_n 's, if c_n increases without limit.

6. Expressions Bounding the Truncation Error

Our approach to the problem of bounding the truncation error makes use of the Schwarz inequality,

$$\left(\sum_{n=0}^{\infty} f_n g_n \right)^2 \leq \left(\sum_{n=0}^{\infty} |f_n|^2 \right) \left(\sum_{n=0}^{\infty} |g_n|^2 \right). \quad (15)$$

We write the truncated approximation to (2) in the form

$$F_N(z) = \sum_{n=0}^{N-1} (c_n F_n) \{c_n^{-1} U_n^k(z)\}. \quad (16)$$

Let us also define two infinite sequences, one for coefficients and the other for polynomials. The rules which we follow are as follows:

$$\bar{F}_n = \begin{cases} F_n, & n < N, \\ 0, & n \geq N, \end{cases} \quad (17)$$

$$\bar{U}_n^k = \begin{cases} U_n^k, & n < N, \\ 0, & n \geq N. \end{cases} \quad (17')$$

We can then write the truncated series in alternative forms,

$$F_N(z) = \sum_{n=0}^{\infty} (c_n \bar{F}_n) \{c_n^{-1} U_n^k(z)\}, \quad (18)$$

$$F_N(z) = \sum_{n=0}^{\infty} (c_n F_n) \{c_n^{-1} \bar{U}_n^k(z)\}. \quad (18')$$

If we subtract (18) and (18') from (2'), and apply the Schwarz inequality, we obtain the expressions we seek,

$$|F - F_N| \leq \|F - F_N\|_r D^k(z), \quad (19)$$

and

$$|F - F_N| \leq \|F\|_r D_N^k(z), \quad (19')$$

where in agreement with eq (9),

$$\{D_N^k(z)\}^2 = \sum_{n=N}^{\infty} \{c_n^{-1} U_n^k(z)\}^2. \quad (20)$$

One should note here that (19) and (19') are by no means the only such inequalities which can be written. One can modify (17) and (17') in many ways, each leading to an inequality. But the expressions written above represent limiting cases of greater interest than some of the other options.^{2a}

By and large, we would expect (19) to be of greatest usefulness when the function being approximated is exactly known, perhaps through an expression which gives the F_n for all values of n . Under these circumstances, the norm of the difference function, $\|F - F_N\|_r$, can be determined with high accuracy and certainty, and simple expressions for an upper bound can be derived.

Alternatively, we would expect (19') to be more useful when the function being approximated is known only in regards to some very general properties, such as asymptotic trends or discontinuous or singular behavior. Under these circumstances the norm of the difference function cannot be accurately known. One may have only some guidance on expected trends with large values of n of the coefficients F_n together with the sequence of estimates of the norm provided by the terms up to the index of truncation. Even though this information is incomplete, it still permits *estimates* of error bounds for all values of z through an estimation of the likely magnitude of a single number. This information is useful, the more so in combination with the sequence of approximations obtained by changing the index of truncation.

The case (19') is of greater use in the calculation of gamma ray and neutron distributions, in which we have a particular interest.

^{2a} Note that D_N^k can replace D^k in (19) to give a sharper form.

In a general way, one can say that the c_n 's provide a mechanism for translating some very general information about the function $F(z)$ into a reduction of the error bounds of a truncated approximation. If one knows that the F_n 's decrease strongly with large n —and this information would follow from location and type of singularities of the Fourier transform—then one can use c_n values which increase rather strongly with increasing values of n . A proper choice would result in only a small increase in either $\|F - F_N\|_r$ or $\|F\|_r$, while it would decrease D^k and D_N^k substantially for all values of z . The result is therefore a narrowing of the estimated bounds for the truncation error.

6.1. Error Limit Functions

Table 1 presents examples of “error limit functions.” The second column gives D^k , and the remaining columns give ratios D_N^k/D^k . These data, somewhat augmented, are also presented in figure 1. For this case and for others to be mentioned, the c_n have the form

$$c_n = (n+1)^{j/4}, \quad (21)$$

and in table 1, and figure 1, the results are for $j=2$.

Note that there are N minima in figure 1, and that they are spaced more widely apart at larger z . Note also that from one value of N to the next a drop by perhaps a factor of $\sqrt{2}$ occurs, while the final rise to unity advances by about $\Delta z=4$ to the right.

TABLE 1. Values for the functions $D^0(z)$ and D_N^0/D^0 given by eqs (9) and (20), using the expression (21) for c_n with $j=2$

z	D^0	$N=1$	2	4	7
0.00	1.124	0.4561	0.3302	0.2293	0.1643
.25	1.048	.2997	.1607	.06876	.02627
.50	1.019	.1903	.07817	.05558	.05302
.75	1.009	.1334	.1007	.09329	.06760
1.00	1.011	.1492	.1492	.1094	.06037
1.50	1.040	.2735	.2142	.09805	.02875
2.00	1.091	.3993	.2332	.06267	.02468
3.00	1.246	.5970	.1860	.06151	.03530
4.00	1.470	.7331	.1308	.09293	.01661
5.00	1.770	.8252	.2071	.07939	.02489
6.00	2.163	.8867	.3439	.05320	.02746
7.00	2.671	.9273	.4789	.07480	.01733
8.00	3.328	.9538	.5971	.1147	.01821
10.00	5.264	.9818	.7736	.1369	.02865
12.00	8.488	.9930	.8810	.1020	.01864
14.00	13.88	.9974	.9408	.1970	.03812
16.00	22.94	.9990	.9719	.3750	.03721
18.00	38.24	.9997	.9872	.5499	.03231
20.00	64.19	.9999	.9944	.6953	.06604
22.00	108.4	1.000	.9976	.8047	.08923
25.00	239.8	1.000	.9994	.9084	.07327
30.00	917.8	1.000	.9999	.9790	.2758

Figure 2 shows the effect of changing the value of j in (21). When j is increased by unity, the curves are reduced by a factor which is much larger at small z than large. The positions of the minima are essentially unchanged.

Figure 3 illustrates changes when k is increased. The main effects are as follows: (a) there is a shift of all minima to the right, this shift being larger for large z than for small z ; and (b) there is a substantial increase in D_N^k/D^k , particularly for values of z near zero. This increase reflects the reduction of the small z region's contribution to the polynomial coefficients, due to the factor

of z added to the weight function. Beyond the small z region, equal increases in k and j have nearly compensating effects on the ordinates of corresponding maxima and minima.

Figure 4 shows a number of functions $D^k(z)$, for different values of k and j . The trend with z does not change very much, although the curves are reduced at large z when j is increased. Increases in the value of k show a correspondingly stronger divergence for $z \rightarrow 0$. Beyond this one might say that the main effect of unit increase in k is a shift of the curve to the right by about $\Delta z = 1$.

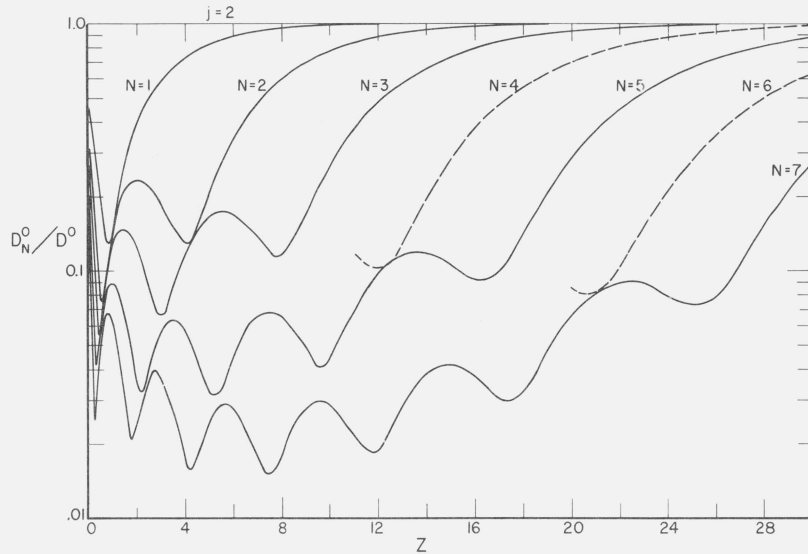


FIGURE 1. Data from table 1, for functions defined by eqs (9) and (20), with $c_n = \sqrt{n+1}$ and $k=0$.

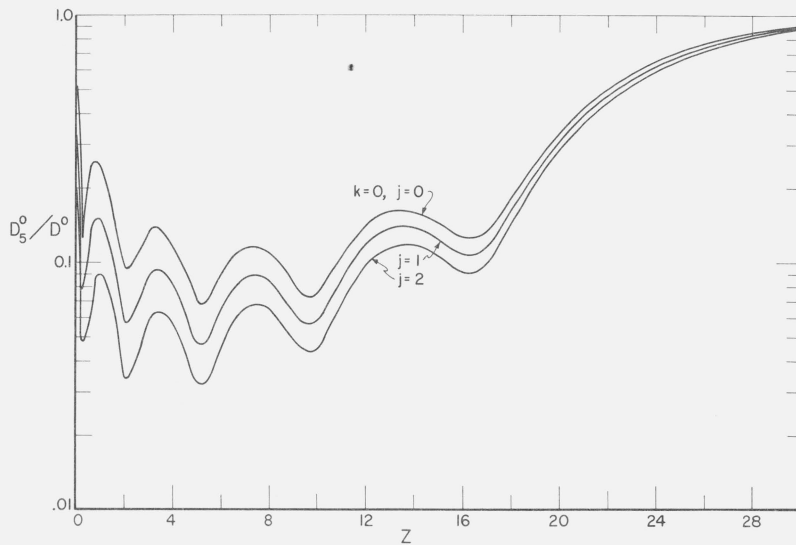


FIGURE 2. The ratios D_N^0/D^0 , for $N=5$, and $c_n = (\sqrt[n]{n+1})^j$.

Data for different values of j are compared.

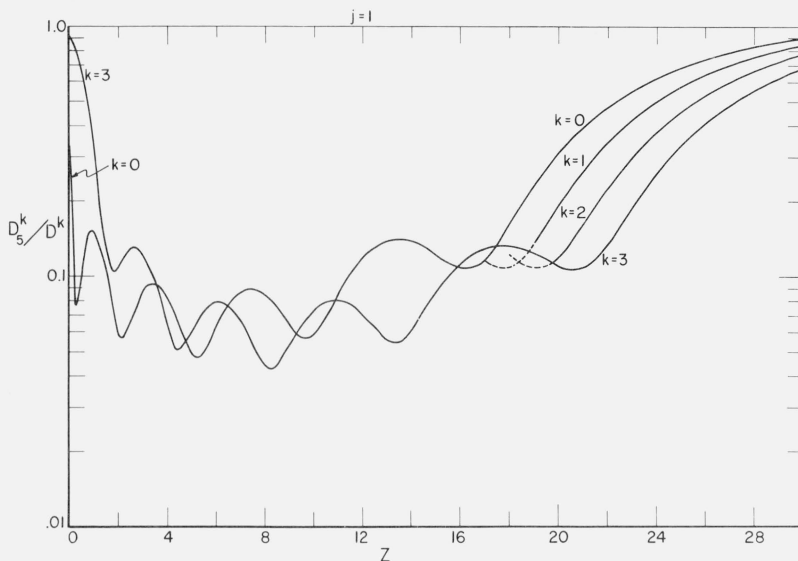


FIGURE 3. The ratios D_N^k/D^k , for $N=5$ and $c_n = (\sqrt[4]{n+1})^{k+1}$.

Data for different values of k are compared.

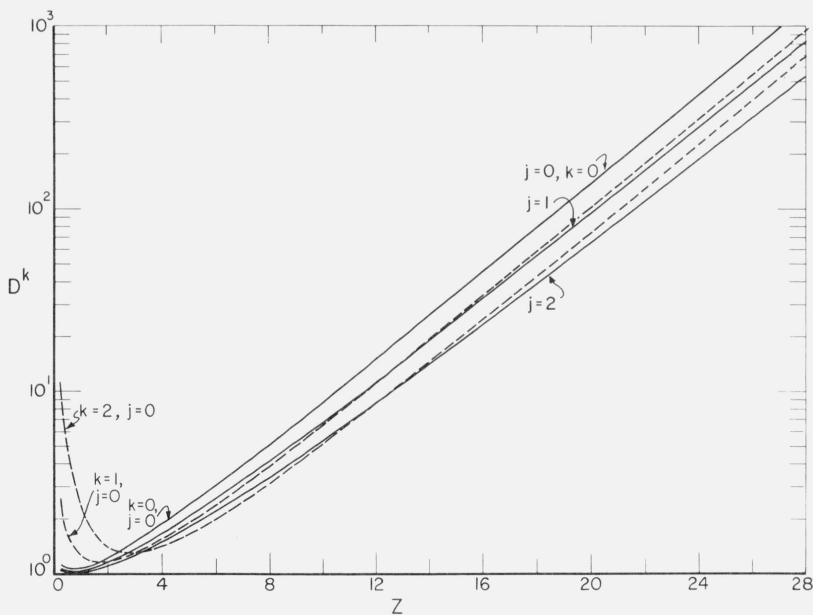


FIGURE 4. The functions $D^k(z)$, for several values of k , using $c_n = (\sqrt[4]{n+1})^j$, with j also taking different values.

6.2. Fractional Error Limits

Using (19) and (19'), we write the ratio of error limits to function $F(z)$ being approximated as follows,

$$\frac{|F - F_N|}{F(z)} \leq Eb(z) \left\{ \frac{\|F - F_N\|_r}{\|F\|_r} \right\}, \quad (22)$$

$$\leq Eb(z) \left\{ \frac{D_N^k(z)}{D^k(z)} \right\} \quad (22')$$

where $Eb(z)$ will be referred to as an "error buildup" function and is given by

$$Eb(z) = D^k(z) \frac{\|F\|_r}{F(z)}. \quad (23)$$

The ratio D_N^k/D^k is shown in figures 1-3, and depends on the approximation system but not on the function being approximated. Further, its general characteristics are largely determined by the number of moments used and other input information. On the other hand, $Eb(z)$ depends on the function $F(z)$ and can be modified by making changes in this function of a manipulative type, which do not necessarily represent additional input information. Further, if $F(z)$ cannot be well-represented, the error buildup function exhibits the situation clearly.

Thus the representation (22') divides the error limits into system-dependent and function-dependent parts. (This is not true if use of (19) rather than (19') results in replacing D_N^k/D^k by $\|F - F_N\|_r/\|F\|_r$.)

Clearly the fractional error limits cannot be made small unless $Eb(z)$ is kept within a reasonable range of values. The trend of $D^k(z)$ shown in figure 4 is strongly increasing, hence $F(z)$ is ideally also a rather strongly increasing function of z .

To illustrate these ideas let us consider the simple case $F(z) = 1$. Since the norm of this function is unity for all values of k , $Eb(z)$ is given by $D^k(z)$. From figure 4, using $k=0$, $j=0$, one thus obtains a value of 8.6 for Eb at $z=10$, and from calculations similar to those of table 1 we evaluate D_N^0/D at 0.051, so that the fractional error limits are about 8.6×0.051 , or about ± 44 percent. On the other hand, if one uses (19) and (23), since $\|F - F_N\|$ is known to be zero for $N \geq 1$, the error limits vanish.

For contrast we consider next the simple function $F(z) = e^{z/5}/(1.25)^{k+1}$, again with $k=0$, and $j=0$. From eq (4) one sees that $F_n = (-9/16)^n$, hence $\|F\|_r = \sqrt{1.46}$; and $\|F - F_N\|_r$, with $N=7$, has the value $\sqrt{1.46} \cdot 0.0178$. The function $Eb(z)$ turns out to be less than three out to a value of z at least as great as $z=16$, and it is no greater than about 6 out to $z=25$. (These comments omit the logarithmic divergence *very* close to $z=0$).

These results enable us to determine that according to eq (19), the error limits are below 5 percent out to $z=16$, and below about 10 percent out to $z=25$. On the other hand, by use of (19'), and data for the $j=0$ case similar to that of table 1 and figure 1, we determine that error limits fluctuate below about 15 percent out to about $z=16$, but rise to about 50 percent at $z=25$.

In a general way, one might say that the narrower limits given by (19) are due to use of more refined information about $F(z)$ than is required by (19').

6.3. Scale Factors

The two functions $F(z) = 1$, and $F(z) = e^{z/5}/(1.25)^{k+1}$, can be regarded as differing mainly in regards to scale factor. If we multiply both by e^{-z} we see this more clearly. The preceding discussion suggests that the error buildup function can be modified to advantage by changes in scale factor which are very easy to make.

To state this in general form, we depart from the practice to this point of using only e^{-z} as weight factor. More generally, we represent a function $F(z)$ with the approximation

$$F(z) \approx \sum_{n=0}^N c_n F_n \left\{ \frac{1}{c_n} U_n^k(z/\xi) \right\}, \quad (24)$$

where the constants c_n have the interpretation previously given, and the scale factor ξ is arbitrary. Expansion coefficients F_n are given by

$$\int_0^\infty d(z/\xi) \hat{U}_n^k(z/\xi) (z/\xi)^k e^{-z/\xi} F(z). \quad (25)$$

If the distribution $F(z)$ has an asymptotic trend which is exponential, with coefficient $(\mu_m - 1/\xi) \neq 0$,

$$F(z) \rightarrow f(z)e^{-\mu_m z + z/\xi},$$

where f is some function which is weaker than exponential, then a finite norm requires that $\frac{1}{\xi} < \mu_m \cdot \sqrt{2}$. Beyond this, however, the choice of ξ is a practical matter, affected by rate of convergence and the resulting size of $Eb(z)$. It turns out that $Eb(z)$ is apt to be better behaved if ξ is chosen a little smaller than $1/\mu_m$, say $\xi \sim \frac{.8}{\mu_m}$. But other considerations can sharpen the selection of scale factor considerably.

6.4. Trends Near Zero

It may happen that the trend of $F(z)$ near zero is known precisely. One can then approximate a modified function, $F(z) - C(z)$, where $C(z)$ has the known trend for $z \rightarrow 0$, tends to zero more strongly than $F(z)$ for $z \rightarrow \infty$, but is otherwise arbitrary. The combination $F(z) - C(z)$ can then be proportional to z or z^2 for small z , so that approximation by U_n^1 or U_n^2 polynomials is appropriate, rather than U_n^0 polynomials.

This situation has many advantages. The U_n^k systems for higher k values are not so accurate for small z , but extend their accuracy considerably farther in the direction of large z values. But if the small values of z cause no problem, due to use of known trends, the accuracy at larger z values is obtained essentially without cost. Use of this additional information thus turns out to give greater accuracy for all z values, small as well as large, as well as narrower bounds on the error.

If, in addition, many moments of the unknown function are available for use, it is possible to omit the lowest moments, while at the same time increasing the value of k even more. This decreases the accuracy for small z values somewhat, but in such a way that the decrease in both accuracy, and knowledge of the accuracy, is not necessarily large enough to be important. At the same time, the approximation for intermediate and large z may be considerably improved. This suggestion of omitting low moments and using large k values was made by E. Morris [13]. Morris did not attempt to obtain a single approximation, but used this approach to extend a low k approximation to much larger z values.

Example 1: The Sievert Integral. Standard tabulations exist of the Sievert integral [14],

$$S(z, \theta) = \int_0^\theta d\phi e^{-z/\cos \phi}, \quad (25)$$

but due to the fact that two variables are involved, additional representations may well prove useful. Hence the evaluation in terms of U_n polynomials is perhaps a nontrivial example of the use of error functions of the type (20'). Writing

$$\cos \theta = \sqrt{1 - \beta}$$

and expanding by means of eq (4), we evaluate the integral term by term to obtain:

$$S(z, \theta) = \sum_{n=0}^{\infty} \frac{(\sin \theta)^{2n+1}}{2n+1} e^{-z} U_n^0(z). \quad (26)$$

If we evaluate this expression, using m terms in the series, the error according to eq (19) is governed by the inequalities

$$(\text{Error})^2 \leq \left\{ \sum_{n=m}^{\infty} \frac{(\sin \theta)^{4n+2}}{(2n+1)^2} \right\} \frac{2}{\pi} K_0(z\sqrt{2}), \quad (27)$$

or, say, using $c_n = (2n+1)^{1/4}$,

$$(\text{Error})^2 \leq \left\{ \sum_{n=m}^{\infty} \frac{(\sin \theta)^{4n+2}}{(2n+1)^{3/2}} \right\} \sum_{n=0}^{\infty} \frac{1}{\sqrt{n+1}} [e^{-z} U_n^0(z)]^2. \quad (27')$$

Note that the error limit given by (27') is finite at $z=0$. Writing ν for $3/2$ or 2 , we have

$$\begin{aligned} \sum_{n=N}^{\infty} \frac{(\sin \theta)^{4n+2}}{(2n+1)^{\nu}} &\leq \frac{(\sin \theta)^{4N+2}}{(2N+1)^{\nu}} \left\{ 1 + \sin^4 \theta \int_{N+1/2}^{\infty} dx \frac{(2N+1)^{\nu}}{(2x)^{\nu}} \right\}, \\ &\leq \frac{(\sin \theta)^{4N+2}}{(2N+1)^{\nu}} \left\{ 1 + \frac{2N+1}{2(\nu-1)} \sin^4 \theta \right\}. \end{aligned} \quad (28)$$

Table 2 gives selected values of the ratio $\|F - F_N\|_r / \|F\|_r$, obtained using (27') and (28). The latter gives an overestimate which for intermediate values of θ can be up to a factor of 2 too high. The norm ratios of table 2 show that (26) gives very accurate values even with only a few terms, throughout a wide range of z values, provided θ is small, say $\theta < \pi/4$.

The approximations of table 3 are given despite being somewhat unimpressive because they illustrate relationships between errors, error limits, scaled and unscaled approximations, and the two approaches to obtaining error limits. For θ near $\pi/2$, eq (26) has slow convergence due to the "sharp" component with $\cos \phi$ near zero. Table 3 gives a variety of types of data for the worst case,

TABLE 2. *Values of $\|F - F_N\|_r / \|F\|_r$ calculated using eq (27')*

The numbers in parenthesis give approximations in which the numerator of the ratio is obtained using eq (28). (Note that the form A-B means $A \times 10^{-B}$ in the top entries.)

θ	$N=2$	4	6
0	0	0	0
15°	1.3-3 (1.4-3)	0.39-5 (0.39-5)	1.3-8 (1.4-8)
30°	1.9-2 (2.1-2)	.77-3 (.94-3)	0.36-4 (.48-4)
45°	0.079 (.11)	.013 (.021)	.0025 (.0046)
60°	.20 (.31)	.075 (.16)	.033 (.070)
75°	.36 (.52)	.23 (.37)	.16 (.29)
90°	.54 (.57)	.46 (.47)	.41 (.42)

TABLE 3. *Fractional error and error limit data for $S(z, \pi/2)$*

To obtain the quantities in parenthesis, a scaled function was calculated, with $\xi=0.8$. This reduces the Eb function strongly at large z . N is the number of terms in the approximating series. Data in the "error" column is fractional. Fractional error limits can be obtained by multiplication of Eb by D_N^0/D^0 .

z	Eb	D_1^0/D^0	$N=4$ Error	D_7^0/D^0	$N=7$ Error
0	0.97 (1.16)	0.366 (0.366)	-0.18 (-0.15)	0.287 (0.287)	-0.14 (-0.11)
0.25	.96 (1.08)	.106 (.079)	-.014 (-.003)	.046 (.047)	.008 (.014)
.50	1.02 (1.11)	.106 (.137)	.04 (.04)	.103 (.120)	.04 (.03)
.75	1.11 (1.16)	.158 (.173)	.06 (.04)	.123 (.112)	.04 (.02)
1.00	1.22 (1.21)	.175 (.168)	.07 (.04)	.105 (.076)	.03 (.01)
1.5	1.48 (1.31)	.148 (.108)	.05 (.01)	.050 (.041)	.002 (-.01)
2	1.76 (1.40)	.095 (.070)	.01 (-.01)	.045 (.060)	-.02 (-.01)
4	3.17 (1.79)	.128 (.106)	-.08 (-.02)	.027 (.038)	.007 (.006)
7	8.09 (2.61)	.098 (.164)	.17 (.05)	.025 (.036)	-.004 (-.008)
10	19.4 (3.69)	.163 (.130)	.29 (-.04)	.038 (.030)	-.07 (.01)
15	91.2 (6.79)	.313 (.640)	4.0 (-.42)	.051 (.052)	.35 (-.03)

$\theta = \pi/2$. Two sets of data, corresponding to 4-term and 7-term approximations, are shown. In all calculations, $k=0$ and $j=1$. Data (in parenthesis) are also given for another calculation in which a scaling of $\xi=0.8$ was used to reduce both the error and the error limits. The two factors giving error limits, $Eb(z)$ and D_N^0/D^0 , are presented separately.

Comparison of the data in parenthesis for Eb with that not in parenthesis demonstrates the reduction of this factor which the scaling brings about. The other factor, D_N^0/D^0 , is not greatly affected by the scaling except that the eventual rise to unity is moved to smaller z value, as already noted.

Errors given for the scaled approximation are almost uniformly smaller, and this is especially marked for large z values.

Error limits calculated by multiplying Eb by D_N^0/D^0 are usually two or more times larger than the actual errors. The pattern of error sizes shows a marked tendency to follow the pattern of error-limit sizes. This is expected, because both are affected by the maxima and minima of the first few polynomials not included in the approximation.

Example 2: Point and Plane Isotropic Gamma Ray Distributions. In the calculation of gamma ray distributions, one can classify components of the distribution by an index which gives the number of scattering interactions that the photons have had with atoms of the surrounding material. Thus one speaks of unscattered, once-scattered, twice-scattered, etc., components. The multiply-scattered component is the sum of all components scattered at least twice.

The advantage of this classification system is in part due to the fact that unscattered and once-scattered components can be described by analytic expressions for many configurations. The multiply-scattered component represents the primary objective of the polynomial representation referred to in this paper; and for a configuration in which the photon source is idealized as a point, the multiply-scattered component is known to vary approximately as z^2 near the source.

One procedure for accurate calculation of distributions is thus to determine unscattered and once-scattered components by direct functional evaluation and quadrature, and to evaluate the multiply-scattered component by polynomial representation. Due to the known z^2e^{-z} dominance in the spatial trend of this component, we have used U_N^2 polynomials. It turns out that the z -dependence of the function multiplying z^2e^{-z} , which is our $F(z)$ to be determined, has a trend which produces an error buildup ratio (22') which does not increase satisfactorily. Hence we have used a value $\xi=1/1.2$ to decrease the error limits. Because the norm is not known, but the convergence of the norm appears to be very good, we have used the largest of the sequence of norm values shown in table 4, together with an error limit expression of the type (19') with $c_m = \sqrt[m]{m+1}$, i.e., $j=1$, to give the values in the next to last error column of table 4.³

Table 4 shows unscattered and once-scattered components as well as the multiply-scattered component. The sum of all components is given in the "Total" column. The last column on the right gives an indication of the truncation error, through use of the difference between approximations using 6 and 8 moments. Both error columns are divided by "Total", rather than by the multiply-scattered result. This explains the tendency for errors to vanish as $z \rightarrow 0$, since the "Total" contains the unscattered component, which is finite in this limit.

The data of table 5 were obtained from the data of table 4 by numerical integration, using

$$I_{PLI}(z) = \frac{1}{2} \int_{|z|}^{\infty} \frac{dr}{r} I_{PTI}(r), \quad (30)$$

³ We use a smaller value of j than (14) would suggest because a factor z^2 is applied to the error limits here. Resulting convergence to zero shown on the table at $z=0$ is conjectured, rather than proven.

TABLE 4. *Gamma ray energy deposition from a point isotropic, Cs-137 source in air*

The first two columns give distance in g/cm² and in feet. The next three columns give unscattered, once scattered, and multiply scattered components, respectively. The column labeled "Estimated Err Lim" gives an estimated fractional error limit to the "Total" column. The final column gives the fractional contribution of the last two terms in the series. For this calculation a total of 8 terms was used in the expansion. Integration of the total with respect to distance (in g/cm²) gives unity.

Energy deposition in air								
g/cm ²	Feet	0 Scat	1 Scat	2 + Scat	Scat/Unscat	Total	Estimated Err Lim	Final Two Terms
0.0000	0.00	2.943-02	0.000	0.000	0.000	2.943-02	0.000	0.000
.1102+00	3.00	2.919-02	2.189-04	2.049-06	7.571-03	2.941-02	7.610-04	4.147-06
.1205+00	3.28	2.916-02	2.389-04	2.447-06	8.274-03	2.940-02	9.063-04	4.946-06
.1469+00	4.00	2.910-02	2.898-04	3.628-06	1.008-02	2.940-02	1.335-03	7.314-06
.2204+00	6.00	2.894-02	4.288-04	8.098-06	1.510-02	2.938-02	2.925-03	1.620-05
.2938+00	8.00	2.878-02	5.643-04	1.428-05	2.011-02	2.935-02	5.063-03	2.834-05
.3673+00	10.00	2.861-02	6.967-04	2.214-05	2.512-02	2.933-02	7.702-03	4.358-05
.4407+00	12.00	2.845-02	8.262-04	3.162-05	3.015-02	2.931-02	1.080-02	6.176-05
.5509+00	15.00	2.821-02	1.015-03	4.881-05	3.771-02	2.927-02	1.620-02	9.420-05
.7346+00	20.00	2.781-02	1.317-03	8.506-05	5.041-02	2.922-02	2.693-02	1.609-04
.9182+00	25.00	2.742-02	1.604-03	1.303-04	6.324-02	2.916-02	3.931-02	2.414-04
.1102+01	30.00	2.704-02	1.877-03	1.839-04	7.622-02	2.910-02	5.285-02	3.338-04
.1205+01	32.80	2.682-02	2.024-03	2.174-04	8.355-02	2.907-02	6.079-02	3.899-04
.1469+01	40.00	2.628-02	2.384-03	3.141-04	1.027-01	2.898-02	8.180-02	5.466-04
.2204+01	60.00	2.484-02	3.265-03	6.529-04	1.577-01	2.875-02	1.385-01	1.040-03
.2938+01	80.00	2.347-02	3.993-03	1.073-03	2.158-01	2.853-02	1.834-01	1.556-03
.3673+01	100.00	2.218-02	4.593-03	1.550-03	2.770-01	2.832-02	2.109-01	2.036-03
.4407+01	120.00	2.096-02	5.085-03	2.066-03	3.412-01	2.811-02	2.206-01	2.441-03
.5509+01	150.00	1.925-02	5.652-03	2.876-03	4.430-01	2.778-02	2.076-01	2.860-03
.7346+01	200.00	1.671-02	6.238-03	4.230-03	6.265-01	2.718-02	1.491-01	3.013-03
.9182+01	250.00	1.450-02	6.493-03	5.486-03	8.260-01	2.648-02	1.156-01	2.564-03
.1102+02	300.00	1.259-02	6.518-03	6.579-03	1.040+00	2.568-02	1.428-01	1.696-03
.1205+02	328.00	1.163-02	6.457-03	7.107-03	1.166+00	2.519-02	1.650-01	1.103-03
.1469+02	400.00	9.485-03	6.136-03	8.184-03	1.510+00	2.381-02	2.023-01	-5.241-04
.2204+02	600.00	5.384-03	4.701-03	9.310-03	2.602+00	1.940-02	1.689-01	-3.517-03
.2938+02	800.00	3.056-03	3.265-03	8.649-03	3.898+00	1.497-02	1.002-01	-2.931-03
.3673+02	1000.00	1.735-03	2.152-03	7.221-03	5.403+00	1.111-02	1.182-01	-3.546-04
.4407+02	1200.00	9.849-04	1.374-03	5.635-03	7.117+00	7.994-03	1.193-01	2.080-03
.5509+02	1500.00	4.212-04	6.746-04	3.566-03	1.007+01	4.662-03	6.924-02	3.141-03
.7346+02	2000.00	1.023-04	1.936-04	1.434-03	1.591+01	1.730-03	9.283-02	-6.928-04
.9182+02	2500.00	2.483-05	5.308-05	5.154-04	2.290+01	5.933-04	5.801-02	-3.566-03
.1102+03	3000.00	6.028-06	1.415-05	1.733-04	3.110+01	1.935-04	9.285-02	-1.135-03
.1469+03	4000.00	3.553-07	9.570-07	1.731-05	5.140+01	1.862-05	9.281-02	5.780-03
.1807+03	4920.00	2.627-08	7.747-08	1.875-06	7.432+01	1.979-06	1.419-01	-3.699-03
.1939+03	5280.00	9.480-09	2.879-08	7.703-07	8.429+01	8.086-07	1.288-01	-9.792-03
.2327+03	6336.00	4.769-10	1.557-09	5.439-08	1.173+02	5.642-08	3.542-01	-1.410-02
.2909+03	7920.00	5.380-12	1.904-11	9.471-10	1.796+02	9.715-10	8.743-01	6.110-02
.3394+03	9240.00	1.282-13	4.775-13	3.001-11	2.378+02	3.061-11	1.860+00	1.716-01
Norm sequence:	0.1929-01	0.2001-01	0.2041-01	0.2043-01	0.2045-01	0.2045-01	0.2046-01	0.2046-01

where I_{PLI} and I_{PTI} are point and plane isotropic source distributions respectively.⁴ This integration was also performed using data for error limits and the difference between approximations; in this way the results shown in the error columns of table 5 were obtained.⁵

⁴ This elementary integral occurs when the plane isotropic source is considered to be a uniform distribution of point isotropic sources over the plane.

⁵ Note that if the error in I_{PTI} is designated by δI_{PTI} while the error limit is L , with $L \geq 0$, then

$$\int \frac{dr}{r} L \geq \left| \int \frac{dr}{r} \delta I_{PTI} \right|;$$

and integration of L thus gives a limit for $|\delta I_{PLI}|$.

TABLE 5. *Gamma ray energy deposition from a plane isotropic Cs-137 source in air.*

[Interpretation of the columns is the same as in table 4.]

Energy deposition in air								
g/cm ²	Feet	0 Scat	1 Scat	2+ Scat	Scat/Unscat	Total	Estimated Err Lim	Final Two Terms
0.1102+00	3.00	6.181-02	9.241-03	9.142-03	2.974-01	8.019-02	1.210-01	7.170-04
.1205+00	3.28	6.051-02	9.230-03	9.142-03	3.036-01	7.888-02	1.230-01	7.288-04
.1469+00	4.00	5.762-02	9.204-03	9.142-03	3.184-01	7.596-02	1.277-01	7.565-04
.2204+00	6.00	5.173-02	9.132-03	9.140-03	3.532-01	7.000-02	1.383-01	8.197-04
.2938+00	8.00	4.758-02	9.061-03	9.139-03	3.825-01	6.578-02	1.468-01	8.705-04
.3673+00	10.00	4.438-02	8.991-03	9.137-03	4.085-01	6.251-02	1.541-01	9.137-04
.4407+00	12.00	4.178-02	8.922-03	9.134-03	4.322-01	5.983-02	1.605-01	9.515-04
.5509+00	15.00	3.862-02	8.820-03	9.130-03	4.648-01	5.657-02	1.687-01	1.001-03
.7346+00	20.00	3.459-02	8.653-03	9.121-03	5.139-01	5.236-02	1.801-01	1.068-03
.9182+00	25.00	3.150-02	8.490-03	9.109-03	5.586-01	4.910-02	1.892-01	1.122-03
.1102+01	30.00	2.902-02	8.332-03	9.095-03	6.005-01	4.645-02	1.966-01	1.165-03
.1205+01	32.80	2.782-02	8.245-03	9.086-03	6.230-01	4.515-02	2.002-01	1.185-03
.1469+01	40.00	2.518-02	8.026-03	9.060-03	6.785-01	4.227-02	2.075-01	1.225-03
.2204+01	60.00	1.999-02	7.457-03	8.965-03	8.214-01	3.642-02	2.182-01	1.260-03
.2938+01	80.00	1.651-02	6.937-03	8.843-03	9.555-01	3.229-02	2.194-01	1.208-03
.3673+01	100.00	1.397-02	6.457-03	8.698-03	1.085+00	2.912-02	2.153-01	1.087-03
.4407+01	120.00	1.200-02	6.016-03	8.535-03	1.213+00	2.655-02	2.089-01	9.105-04
.5509+01	150.00	9.754-03	5.417-03	8.261-03	1.402+00	2.343-02	1.993-01	5.725-04
.7346+01	200.00	7.163-03	4.559-03	7.754-03	1.719+00	1.948-02	1.924-01	-9.815-05
.9182+01	250.00	5.421-03	3.847-03	7.213-03	2.040+00	1.648-02	1.969-01	-7.855-04
.1102+02	300.00	4.185-03	3.252-03	6.663-03	2.369+00	1.410-02	2.025-01	-1.393-03
.1205+02	328.00	3.645-03	2.963-03	6.358-03	2.557+00	1.297-02	2.027-01	-1.676-03
.1469+02	400.00	2.597-03	2.339-03	5.599-03	3.056+00	1.054-02	1.941-02	-2.147-03
.2204+02	600.00	1.101-03	1.221-03	3.792-03	4.553+00	6.114-03	1.498-01	-1.720-03
.2938+02	800.00	5.015-04	6.466-04	2.485-03	6.245+00	3.633-03	1.338-01	1.653-04
.3673+02	1000.00	2.384-04	3.449-04	1.593-03	8.129+00	2.176-03	1.308-01	1.857-03
.4407+02	1200.00	1.165-04	1.848-04	1.004-03	1.020+01	1.305-03	1.138-01	2.399-03
.5509+02	1500.00	4.132-05	7.343-05	4.916-04	1.368+01	6.064-04	1.002-01	9.887-04
.7346+02	2000.00	7.820-06	1.599-05	1.433-04	2.037+01	1.671-04	1.006-01	-2.567-03
.9182+02	2500.00	1.557-06	3.536-06	4.042-05	2.823+01	4.551-05	1.007-01	-2.140-03
.1102+03	3000.00	3.206-07	7.903-07	1.118-05	3.733+01	1.229-05	1.244-01	1.954-03
.1469+03	4000.00	1.451-08	4.046-08	8.211-07	5.939+01	8.761-07	1.481-01	3.143-03
.1807+03	4920.00	8.839-10	2.678-09	7.140-08	8.380+01	7.496-08	2.021-01	-1.130-02
.1939+03	5280.00	2.985-10	9.293-10	2.726-08	9.443+01	2.848-08	2.465-01	-1.634-02
.2327+03	6336.00	1.263-11	4.205-11	1.596-09	1.297+02	1.651-09	5.964-01	-6.166-04
.2909+03	7920.00	1.152-13	4.134-13	2.197-11	1.943+02	2.250-11	1.415+00	1.177-01
.3394+03	9240.00	2.365-15	8.910-15	5.840-13	2.507+02	5.953-13	3.300+00	2.583-01

The good quality of the results shown in tables 4 and 5 is due largely to the predominance of the exponential trend of the unscattered photons throughout all scattered components. The approximations would not be satisfactory otherwise.

One should be aware that the error limits presented in tables 4 and 5 only pertain to one type of error. They do not include errors due to round-off in the calculation of moments, errors due to the finite numerical integration mesh in the calculation of moments or in the quadrature over the once-scattered components. And they take no cognizance of limits to the accuracy of the assumed interaction cross sections.

7. References

- [1] Spencer, L. V., and Fano, U., Penetration and diffusion of x-rays. Calculation of spatial distributions by polynomial expansion, J. Res. Nat. Bur. Stand. (U.S.) **46**, 446-456, (June 1951) RP2213.
- [2] Fano, U., Spencer, L. V., and Berger, M. J., Penetration and diffusion of x-rays, Handbuch der Physik V. **38** (Springer-Verlag, Berlin, 1959).
- [3] Stanley, P., An investigation of biorthogonal polynomials derivable from ordinary differential equations of the third order, p. 38, J. Math. Anal. Appl. **4**, No. 1 (Feb. 1962).
- [4] Spencer, L. V., and Flusser, P., Properties of a useful biorthogonal system, J. Res. Nat. Bur. Stand. (U.S.), **71B** (Math. and Math. Phys.), No. 4, 197-211 (Oct.-Dec. 1967).
- [5] Goldstein, H., and Wilkins, J. E., Jr., Calculation of the Penetration of Gamma rays (Nucl. Development Associates, Inc., NYO-3075, 1954).
- [6] Speilberg, D., Penetration of Neutrons from a Point Fission Source in Air—Moments Method of Calculation (Nucl. Development Associates, Inc., Report NDA 2106-10, April 28, 1961).
- [7] Spencer, L. V., Structure shielding against fallout radiation from nuclear weapons, Nat. Bur. Stand. (U.S.), Monogr. **42**, 136 pages (June 1942).
- [8] Shimizu, A., Calculation of the penetration of gamma rays through slabs by the method of invariant imbedding, Nucl. Sci. & Engr. **32**, 184 (1968).
- [9] Raso, D. J., Monte Carlo calculations on the reflection and transmission of scattered gamma rays, Nucl. Sci. & Engr. **17**, 414 (1963).
- [10] Doggett, W. O., and Bryan, F. A., Jr., Theoretical dose transmission and reflection probabilities for 0.2-10.0 MeV photons obliquely incident on finite concrete barriers, Nucl. Sci. & Engr. **39**, 103 (1970).
- [11] Spencer, L. V., and Lamkin, J., Slant penetration of γ -rays in H_2O , unpublished report of NBS (July 17, 1958), Slant penetration of γ -rays: mixed radiation sources, unpublished report (Feb. 27, 1959), and Slant penetration of γ -rays in concrete, NBS unpublished report (Nov. 10, 1959).
- [12] Langer, J. S., and Varga, R. S., The Problems of Moments with Application to Neutron Flux Distributions, Westinghouse Electric Corp., WAPD-TN-520 (November 1955).
- [13] Morris, E. E., An extension of the polynomial approximation in moments method calculations to the deep penetration of gamma radiation, Transactions **11**, No. 1, pages 418-419 (1968).
- [14] Abramowitz, M., and Stegun, I. A., Handbook of Mathematical Functions, Nat. Bur. Stand. (U.S.), Appl. Math. Ser. **55**, 1051 pages (June 1964).

(Paper 75B1&2-339)